

Author's Accepted Manuscript

Gene regulation in response to graded hypoxia: The non-redundant roles of the oxygen sensors PHD and FIH in the HIF pathway

Frédéric Dayan, Marc Monticelli, Jacques Pouysségur, Elisabeth Pécou

PII: S0022-5193(09)00116-7
DOI: doi:10.1016/j.jtbi.2009.03.009
Reference: YJTBI5492



www.elsevier.com/locate/jtbi

To appear in: *Journal of Theoretical Biology*

Received date: 24 November 2008
Revised date: 2 March 2009
Accepted date: 5 March 2009

Cite this article as: Frédéric Dayan, Marc Monticelli, Jacques Pouysségur and Elisabeth Pécou, Gene regulation in response to graded hypoxia: The non-redundant roles of the oxygen sensors PHD and FIH in the HIF pathway, *Journal of Theoretical Biology* (2009), doi:[10.1016/j.jtbi.2009.03.009](https://doi.org/10.1016/j.jtbi.2009.03.009)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Gene regulation in response to graded hypoxia: the non-redundant roles of the oxygen sensors PHD and FIH in the HIF pathway

Frédéric Dayan ^{a,b}, Marc Monticelli ^b, Jacques Pouyssegur ^a, Elisabeth Pécou ^b

^a Institut de Signalisation, Biologie du Développement et Cancer - UMR CNRS 6543 - Université de Nice Sophia Antipolis- Centre Antoine Lacassagne - 06000 Nice, France.

^b Laboratoire J.-A. Dieudonné- UMR CNRS 6632 -Université de Nice Sophia Antipolis- Parc Valrose 06104 Nice, France.

Corresponding author: F. Dayan, Frederic.Dayan@unice.fr, phone: 33(0)492076229, Fax: 33(0)493517974

Abstract

HIF, being the master protein involved in adaptation to low $p[O_2]$, plays a major role in many physiological and pathological phenomena: development, inflammation, ischemia and cancer. PHD and FIH are the two oxygen sensors that regulate the HIF pathway. Here we model the regulatory dynamics in an oxygen gradient by a system of differential equations. A part of the work consists in a qualitative analysis, driven independently of the values of the parameters, which explains the non-redundant functional roles of FIH and PHD. In a second part, we use biological experiments to fit the model in a physiologically relevant context and run simulations. Simulation results are confronted with success to independent biological experiments. The combination of biological data and mathematical analysis stresses that FIH is a fine modulator determining whether a given gene should be induced in mildly or in strongly hypoxic areas. Moreover it gives access to other functional predictions that are not directly accessible by pure experiments, for instance the stoichiometry of prolyl-hydroxylation on HIF, and the switch-like properties of the system.

Availability: An interactive simulation interface is available at <http://sdi.ljad.free.fr/spip.php?article111>

Keywords: Gene expression/ HIF/ Mathematical model/ Oxygen sensors/ Switch-like behavior.

1 Introduction

Most mammals live in an atmosphere composed of 21% oxygen, which is considered a physiological environment. Nevertheless there are many circumstances in which a cell can become hypoxic - that is, subjected to low oxygen pressure. For instance an organism can be exposed to extreme conditions such as high altitude. Moreover, hypoxia is a characteristic of many pathologies. Cerebrovascular insult or myocardial infarction are examples of ischemic disease. Oxygen pressure is also strongly reduced in cancer (Semenza, 2003), mainly because of (i) the chaotic architecture of tumor vessels (Folkman, 2007) that leads to lower perfusion, and (ii) the intense proliferation of cancer cells that develop distant to blood vessels.

The common feature of these situations is that hypoxia is a powerful cellular stimulus that induces a signaling pathway dedicated to environmental adaptation (Pouyssegur et al., 2006). In the case of ischemic disease, this adaptation is beneficial to the patient as it promotes survival of cells that are exposed to low oxygen pressure. On the contrary hypoxia is synonymous with bad prognosis in cancer (Semenza, 2003), selecting highly virulent cells that are able to survive in extremely unfavorable environments.

The master protein of the hypoxic pathway is the Hypoxia Inducible Factor (HIF) (for review: (Brahimi-Horn and Pouyssegur, 2007)). HIF is a transcription factor that targets a series of adaptation genes under hypoxic conditions. It is a hetero-dimer composed of a regulatory α subunit and a constitutive β subunit. From a functional view, the potency of this protein to recruit the RNA-polymerase lies in two transactivation domains (TAD) located on HIF α . One of these TAD is C-terminal (named C-TAD) and the other is N-terminal (named N-TAD).

HIF is regulated by oxygen thanks to post-translational modifications of its α subunit. The enzymes that modify HIF α and that are responsible for pO_2 sensitivity are called oxygen sensors. To date, two oxygen sensors have been characterized: HIF-Prolyl-Hydroxylase (PHD) (Ivan et al., 2001); (Jaakkola et al., 2001) and HIF-Asparaginyl-Hydroxylase (also called FIH for Factor Inhibiting HIF) (Lando et al., 2002a), (Lando et al., 2002b); (Mahon et al., 2001). Both are members of the 2-oxoglutarate dependent dioxygenase family and both catalyze a reaction of hydroxylation on HIF α (see Figure 1).

In the presence of oxygen, PHD hydroxylates HIF α on two prolyl residues. As a consequence, HIF interacts with the von Hippel Lindau (*pVHL*) ubiquitin ligase and is targeted for massive proteasomal degradation. Moreover, HIF α is hydroxylated on an asparaginyl residue by the other oxygen sensor FIH. This last modification specifically impairs the C-TAD activity of HIF (For review see (Pouyssegur et al., 2006)). In hypoxic conditions, these two sensors are inactivated by the lack of oxygen substrate, and thus HIF is stabilized and active *via* its two transactivation domains. In such conditions, HIF is able to target a wide series of precise adaptation genes associated with Hypoxia Responsive Elements (HRE) which are specific DNA recognition sequences.

Recently the role of FIH as a pure HIF inhibitor has been questioned (Dayan et al., 2006), (Dayan et al., 2008), (Yan et al., 2007). Importantly, FIH only inhibits the C-TAD activity and not the N-TAD activity. New evidence tend to characterize FIH as a discriminator between a N-TAD and a C-TAD spectrum of genes (Dayan et al., 2006). As a result FIH would qualitatively affect the downstream action of HIF. Here we propose to build a mathematical model of this signaling pathway that integrates these new features, including the differential action of FIH on the two HIF's TAD.

2 Results

2.1 Total HIF protein versus O₂ variation

2.1.1 Analytic expression

Since biological experiments measure the total concentration of HIF proteins, we are interested in the value at equilibrium of the sum

$$[\text{HIF}_{\text{TOT}}] = [\text{HIF}] + [\text{HIFOH}_a] + [\text{HIF}(\text{OH}_p)_2] + [\text{HIFOH}_a(\text{OH}_p)_2],$$

where the brackets $[\cdot]$ denote the concentration.

By a mathematical analysis (see Appendix) the following lemma comes readily:

Lemma 2.1 1. The equilibrium values of $[\text{PHD}]$, $[\text{FIH}]$ and $[\text{VHL}]$ do not depend on oxygen and are equal to:

$$[\text{PHD}]_0 = \frac{k_1 S_1}{l_1} \quad [\text{FIH}]_0 = \frac{k_2 S_2}{l_2} \quad [\text{VHL}]_0 = \frac{k_3 S_3}{l_3}$$

2. Let $[\text{HIF}]_0$ be the reference amount of HIF protein corresponding to anoxia:

$$[\text{HIF}]_0 = \frac{k_0 S_0}{l_0}$$

Let

$$\tau_1([\text{O}_2]) = \frac{a_1 b_1 [\text{PHD}]_0 [\text{O}_2]^2}{d_1 + b_1 [\text{O}_2]^2} \quad (1)$$

$$\tau_3 = \frac{a_3 b_3 [\text{VHL}]_0}{d_3 + b_3} \quad (2)$$

where a_1 , d_1 and b_1 are the kinetic parameters involved in the hydroxylation of HIF by PHD, while a_3 , d_3 and b_3 are the rates of the elementary reactions participating to the VHL-dependent proteasomal degradation of all hydroxylated forms of HIF (see Figure 2). The units of τ_1 and τ_3 are inverse time units t^{-1} .

Then, the total amount of HIF protein at equilibrium depends on oxygen $[\text{O}_2]$ in the following way:

$$[\text{HIF}_{\text{TOT}}] = [\text{HIF}]_0 \frac{1 + \frac{\tau_1([\text{O}_2])}{\tau_3}}{1 + \frac{\tau_1([\text{O}_2])}{l_0}} \quad (3)$$

See Figure 3 for the interpretation of τ_1 and τ_3 .

2.1.2 Analysis of HIF expression

In our mathematical model we have chosen the simplest kinetic law, which is mass-action law. This formalism avoids to make any assumption on the speed of the elementary reactions. However, we noticed afterwards that the parameters τ_i which arise in the formula [3] (derived from the equations of our mass-action model, see Appendix) have a natural interpretation if we apply the quasi-steady state model reduction, as explained thereafter.

In our context, the quasi-steady state model reduction applies to the reactions of hydroxylation of HIF proteins by PHD and FIH and to the interaction of HIF proteins hydroxylated on Pro residues with VHL (see Figure 3). It consists in assuming that the concentration of intermediary complexes HIF proteins-enzyme reach almost immediatly their equilibrium value (the so-called Briggs-Haldane approximation). Then, each hydroxylation (respectively, interaction with VHL) can be considered as one reaction, governed by Michaelis-Menten kinetics (see Section 4.3 for details).

In first approximation, when substrate concentrations (HIF in this case) are low, the enzymatic speed is proportional to the substrate concentration. In this regime,

- τ_1 is the multiplying factor of the rate of hydroxylation by PHD; it depends on the oxygen level;
- τ_3 is the multiplying factor of the rate of degradation through VHL.

Formally, τ_1 and τ_3 are the slopes of the Michealis-Menten curve near the origin. Another qualitative interpretation in terms of enzymatic activity is that τ_1 and τ_3 are equal to V_{\max}/K_M where K_M is attributed to the HIF substrate. The higher this quantity is, the more potent the enzyme is, either in reference to its maximal speed, or to its better affinity for its substrate, HIF. In this context, τ_1 (in reference to PHD) appears in a ratio with τ_3 (in reference to VHL) and in a ratio with l_0 (in reference to the non specific degradation of HIF). If one considers the extreme condition $l_0 = \tau_3$ (which would mean that the non specific degradation of HIF is equally potent to the VHL-dependent degradation), HIF is no longer dependent on oxygen. In that case, the oxygen- and non oxygen-dependent degradation are balanced. If $l_0 > \tau_3$, the function would be increasing. And if $l_0 < \tau_3$ (which is relevant physiologically), the function is decreasing (See Figure 4 (Left)).

The curves in Figure 4 are plotted from a set of parameters adapted from previous publication (Kohn et al., 2004). In this context, which satisfies $\tau_3 > l_0$, HIF is a strictly decreasing function of oxygen. Its sensitivity to PHD is illustrated in the right panel. The sharpness of the HIF response is highly dependent on a_1 , the association constant between HIF and PHD.

2.1.3 Fitting with experiments

In our approach, we choose to determine the parameters of our system that fit the best with experimental data. In the case of HIF, we use a protein quantification by western blotting as a reference. Indeed, Figure 5 (Left) reports a measure of HIF_{TOT} concentration at different pressures of oxygen. We define the function $h([\text{O}_2]) = R[\text{HIF}_{\text{TOT}}]([\text{O}_2])$, where the re-scaling R factor is linked to the fact that quantification is done in an arbitrary unit, and we fit h with the experimental data.

Function h is totally determined by three parameters:

$$a = R[\text{HIF}]_0 \frac{1 + \frac{a_1[\text{PHD}]}{\tau_3}}{1 + \frac{a_1[\text{PHD}]}{l_0}} \quad b = \frac{d_1}{b_1} \frac{1}{1 + \frac{a_1[\text{PHD}]}{\tau_3}} \quad c = \frac{d_1}{b_1} \frac{1}{1 + \frac{a_1[\text{PHD}]}{l_0}}$$

which satisfy the relation:

$$h([\text{O}_2]) = a \frac{b + [\text{O}_2]^2}{c + [\text{O}_2]^2}$$

We evaluated a , b and c by nonlinear fitting with the help of the software *Mathematica*©: Figure 5 (Right).

2.1.4 The endogenous HIF is hydroxylated at both proline sites

In the conception of our model (Section 4.2), we arbitrarily choose to take into account the double hydroxylation of HIF by PHD on Pro 402 and Pro 564 (in addition to the hydroxylation on Asn804 by FIH). Nevertheless, it is still debated if one or two prolines are simultaneously hydroxylated on HIF. It appears that these two reactions are independently driven (Chan et al., 2005). For the first time our model gives access indirectly and easily to the hydroxylation status of the endogenous HIF. Indeed, Table 3 shows an optimization score as a function of the number of prolyl hydroxylations n . This number impacts the enzymatic equations and consequently the profile of HIF_{TOT} within the oxygen gradient. When compared to the experimental data (i.e HIF_{TOT} protein quantification), the optimization calculation characterizes $n = 1.98$ with the best optimization score $d = 2.29$. In contrast to other values of d in Table 3, the double hydroxylation on prolyl is noticeably the best configuration. As a consequence, from the simple protein profile of HIF_{TOT} , and from our mathematical model, we conclude that the endogenous HIF in LS174 cells (colon carcinoma) is hydroxylated on two prolyl residues.

2.1.5 Fitted set of enzymatic parameters

It is clear that for a full set of kinetic parameters $a_i, b_i, d_i, k_i, S_i, l_i$, there corresponds a unique triplet (a, b, c) . Conversely, to each triplet (a, b, c) , there corresponds an infinite set of kinetic parameters. In Table 2, we give an arbitrary set of kinetic parameters corresponding to the best fitted value of a, b , and c .

2.2 Gene stimulation versus O_2 variation

2.2.1 Analytical expression

According to our hypothesis on the roles of both transactivation domains of HIF on HRE stimulation, we define the gene stimulation function by

$$G_q = N + qC \quad (4)$$

where N and C are respectively the N-TAD and C-TAD-related stimulation functions:

$$\begin{aligned} N &= [\text{HIF:HRE}] + [\text{HIFOH}_a\text{:HRE}] + \\ &\quad [\text{HIF}(\text{OH}_p)_2\text{:HRE}] + [\text{HIFOH}_a(\text{OH}_p)_2\text{:HRE}] \\ C &= [\text{HIF:HRE}] + [\text{HIF}(\text{OH}_p)_2\text{:HRE}] \end{aligned}$$

Analogously to τ_1 and τ_3 , we define:

$$\tau_2([O_2]) = \frac{a_2 b_2 [\text{FIH}]_0 [O_2]}{d_2 + b_2 [O_2]} \quad (5)$$

It has inverse time unit and can be interpreted as the multiplying factor of the rate of hydroxylation by FIH in the linear regime of the quasi-steady state reduced model (see Figure 3 and Section 4.3).

In terms of the τ_1, τ_2 and τ_3 , we get (see Appendix):

$$G_q = \frac{a_4}{d_4} [\text{HIF}]_0 \frac{1 + \frac{\tau_1}{\tau_3}}{1 + \frac{\tau_1}{l_0} + \frac{a_4}{d_4} [\text{HIF}]_0 (1 + \frac{\tau_1}{\tau_3})} (1 + q \frac{1 + \frac{\tau_1}{l_0}}{1 + \frac{\tau_1}{l_0} + \frac{\tau_2}{l_0}}) \quad (6)$$

2.2.2 Normalization and numerical fitting

In order to compare the stimulation function between various genes, we introduce a normalized gene stimulation function

$$\begin{cases} g_q = R' G_q \\ g_q(0) = 100 \end{cases} \quad (7)$$

where G_q is defined by Equation [4]. From an experimental view point it can be interpreted as the stimulation level with respect to the stimulation induced in anoxia. We get:

$$g_q = \frac{100(1 + \frac{a_4}{d_4} [\text{HIF}]_0)(1 + \frac{\tau_1}{\tau_3})}{(q + 1)(1 + \frac{\tau_1}{l_0} + \frac{a_4}{d_4} [\text{HIF}]_0(1 + \frac{\tau_1}{\tau_3}))} (1 + q \frac{1 + \frac{\tau_1}{l_0}}{1 + \frac{\tau_1}{l_0} + \frac{\tau_2}{l_0}}) \quad (8)$$

The function g_q was calibrated using data obtained from the gene *ca-IX* (see Appendix for details). A set of enzymatic parameters, compatible with the calibration from experimental data is given in Table 2 and is used in the following sections.

2.3 FIH modulates the profile of HIF-targeted genes

In the context of our mathematical model, we questioned the biological impact of the oxygen sensor FIH. Our aim is to reach a better understanding of the qualitative role of FIH. FIH and PHD are often considered as two redundant lockers of the HIF pathway. Yet in opposition to PHD, FIH was recently revealed not to be a pure inhibitor of the HIF pathway (Dayan et al., 2006). FIH only inhibits the C-TAD-sensitive genes while “N-TAD-dominant” genes (e.g *pgk1*) are not sensitive to this oxygen sensor. As a result FIH would control a switch between the N-TAD and the C-TAD repertoire.

For the first time, our MIM includes this precise level of complexity, and opens the way to monitor the influence of FIH on the N-TAD and C-TAD spectrums. We chose to investigate this effect thanks to an inhibition, Figure 7 (panels D, E, F) and an overexpression of FIH, Figure 7 (panels A,B,C). Note that all parameters in the numerical experiment are fitted to biological data; here the arbitrary value 100 for FIH corresponds to the physiological context. Then, the inhibition of FIH results in a contraction of the gene expression curves. More precisely, curves associated to C-TAD-sensitive genes (highest q) are relocalized into milder hypoxic areas. Simultaneously, N-TAD only genes ($q = 0$) are not affected, which is consistent with the fact that the N-TAD activity is not modified by FIH. At the extreme point where $\text{FIH} = 0$ (complete knockout of

FIH), C-TAD-sensitive gene profiles become superposed to N-TAD-only curves. Consistently, increasing the amount of FIH from 100 to 1000 gives symmetrical results: the overexpression of FIH leads to a relocalization of C-TAD-sensitive genes into stronger hypoxic areas.

The shape of this series of curves reflects the ability of FIH to retain the activity of the C-TAD-sensitive genes, and to release it in strongly hypoxic areas of the oxygen gradient. Conversely, the expression of N-TAD-only genes is not retained by FIH, and should be released in milder hypoxic regions. As a result, FIH is a modulator of the HIF spectrum profile in an oxygen gradient, and allows us to differentiate the gene expression pattern of the N-TAD and C-TAD repertoires of genes.

2.4 Biological predictions supported by experimental data

2.4.1 Prediction of the C-TAD-sensitivity for a series of 25 genes

In this part we exploit experimental data from quantitative PCR performed on a series of 25 HIF-targeted genes. For each gene, a relative induction score of hypoxia is evaluated by calculating the ratio (gene expression at 3 %O₂) / (maximal gene expression under anoxia). From this score, our model attributes a value of q , associated to each gene (Figure 8). Thus our theoretical model gives access to the quantitative contribution of the C-TAD, which is difficult to estimate purely experimentally. See also Appendix.

2.4.2 Modulation of the gene profile, and physiological limitations of the action of FIH

The x -axis in Figure 8 quantifies the flexibility of gene induction given by the action of FIH. This oxygen sensor has the ability to retain the expression of C-TAD-sensitive genes (high q) and to release it only in highly hypoxic areas. Such genes present a low relative induction score in Figure 8 (x -axis), which means that their induction at 3% oxygen is far from their maximal anoxic induction. Thus FIH modulates the profile of C-TAD-sensitive genes; nevertheless, importantly, our model predicts an asymptotic border that defines a limit in the action of FIH: dotted line in Figure 8, y_1 value in Appendix . Whatever the strength of the C-TAD sensitivity (i.e. q), the consequence in terms of gene induction is necessarily limited. In the case of our numerical values, the ratio between the gene induction at 3% oxygen and anoxia should always be superior to $y_1 = 3.5\%$. This a priori counter-intuitive result reflects a physiological limit of flexibility of the gene profile within an oxygen gradient.

Note that in the absence of FIH, another inferior limit of gene induction appears: y_0 value in Appendix , when FIH = 0 (or $q = 0$ for N-TAD-only genes). In that situation, PHD is the only oxygen sensor driving the shape of the gene expression curve from normoxia to anoxia. Consequently, when FIH = 0, the gene expression is parallel to the HIF protein induction, and more limited. In the case of our numerical values, this limit is equal to 80% under the influence of PHD alone. A biological interpretation of y_1 and y_0 relies on the impact of FIH on the gene profile in addition to PHD. Here this enzyme, by playing on the HIF activity, allows to shift the inferior limit from 80% to 3.5%; thanks to FIH, the system is still limited, but substantially more flexible.

2.4.3 The calculated C-TAD sensitivity is consistent with FIH modulation experiments

In order to check the biological relevance of our model and our predictions, we exploited an independent biological experiment. Indeed, as previously published (Dayan et al., 2006), the sensitivity to FIH of a given gene can be estimated thanks to FIH modulation experiments (overexpression and inhibition by shRNA).

Then for each gene, we defined a FIH sensitivity score that is equal to the ratio (gene response following inhibition of FIH)/(gene response following overexpression of FIH). The higher this FIH score is, the more a gene responds to FIH. This score was at the basis of the C-TAD and N-TAD sensitivity classification in (Dayan et al., 2006). Among 25 HIF-targeted genes, 16 were classified as FIH-sensitive / C-TAD responsive and 9 as FIH insensitive / N-TAD dominant. Aiming at putting our model to the test, we compared this FIH sensitivity score to the calculated q predicted by our system. Figure 9 shows that among the same 25 genes, a nice correlation is observed, and that 3 clusters emerge. Cluster A gathers 8 N-TAD dominant genes with low q . Cluster B gathers 10 C-TAD responsive genes with intermediary q . In cluster C, 3 C-TAD responsive genes present a substantially higher q . Thus our model is consistent with the previous classification as N-TAD (green diamonds) and C-TAD-sensitive (blue squares) genes. Moreover, it allows easy discrimination between these 2 categories on the basis of a simple experiment (gene induction at 3% and 0 % oxygen) whereas the previous classification was based on a genetic modulation of FIH.

2.4.4 Exception genes

In Figure 9, 4 genes escape to the 3 clusters that are described above. They are exceptions that don't match with our model. Our interpretation is that their gene expression is driven by other actors that are not present in our model, or still have to be discovered. Interestingly these 4 genes are pointed out thanks to the correlation graph. Coefficient q for *phd3* is exceptionally

high with comparison to its FIH sensitivity score, which reflects that in addition to FIH, another factor should contribute to release *phd3* expression in highly hypoxic zones. On the opposite, *trefoil factor 3* and, to a certain extent, *p21* present a relatively low q when compared to their FIH sensitivity score, which means that an unknown mechanism decreases the effect of FIH on their induction profile within an oxygen gradient. At last, *bnip3* presents an unusual FIH sensitivity score, below 1, which means that FIH does not inhibit *bnip3* in this experiment, and even further that it increases its expression. In our model, such a situation should correlate with a negative value of q (an inhibitory activity of the C-TAD). Nevertheless it is not the case, and certainly the mechanism of *bnip3* induction is not fully described by our MIM.

2.5 Switch-like behavior of the HIF protein

2.5.1 Preliminary definitions.

In the wider acceptance, a system has a switch-like behavior if, by variation of a parameter, it changes from one stable state to another one and the transition is sharp.

Two very distinct phenomena can lie behind an observed switch-like behavior, and distinguishing which of both occurs in a specific situation is biologically relevant.

The first of these phenomena is the switch. A classical example is the lysis/lisogeny switch of the λ -phage (see for instance (Arkin et al., 1998)), but there are plenty of others. A fundamental characteristic of switches is irreversibility: if the change from stable state A to stable state B occurs at a critical value c when increasing the parameter, then if one decreases the parameter below c , the system stays in state B . This property was already pointed out in (Monod and Jacob, 1961) and is now well understood in terms of a hysteresis effect: the parameter c is a bifurcation point and there is another bifurcation point $c' < c$ and for any value of the parameter between c' and c , the system is bistable.

We call a “shift” the case when no irreversibility is observed. Although in both situations a sharp transition is observed between two stable states, the number of stable states is a fundamental qualitative difference between a switch and a shift: while in the case of a switch, both stable states “cohabite” in a region of parameter space, in the case of a shift, whatever the value of the parameter is, the system has a unique stable state.

Let us emphasize that a switch behavior is a robust qualitative property of the system, that is: (1) it does not depend on specific choices of initial conditions and (2) it persists under perturbations. A perturbation can affect the numerical values of the stable states and the bifurcation point but not their existence and the sharp transition. In contrast, the shift property is not robust: it depends on the choice of a sharpness threshold, that is the minimum value of the steepness (the derivative) at the inflexion point. The numerical value for the threshold is empirical and is strongly related to the context. This approach is taken in (Qutub and Popel, 2006).

Finally, notice that many authors assimilate switch and switch-like behavior, but for the purpose of our discussion, we make the careful distinction.

2.5.2 Questioning the switch-like behavior

In (Kohn et al., 2004), a switch-like behavior was numerically observed for the gene stimulation function; and the authors suggested that the underlying phenomenon was a switch. Notice that when we set $\text{FIH} = 0$ and choose the hypothesis of only one hydroxylation of HIF by PHD, our model reduces to a simplified version of the model in (Kohn et al., 2004). In this case the gene stimulation and the total HIF protein functions give qualitatively the same response as in (Kohn et al., 2004).

Following our mathematical analysis (and this discussion is independent on the hypothesis of a unique or a double hydroxylation of HIF by PHD), it appears that the switch-like behavior is strongly dependent on the condition that PHD is not degraded (Remark A.1 in Appendix), a condition which is not in agreement with physiology. To the contrary when PHD is produced and degraded, the switch-like property is lost. To assess this theoretical argument, we simulated the model of Kohn, partially simplified by assuming that the role of the constitutive β subunit of HIF can be neglected in the regulation of hypoxia, with the set of parameters proposed in (Kohn et al., 2004) and with the hypothesis of a unique hydroxylation of HIF by PHD, see Figure 10. Actually, we also simulated the complete Kohn’s model (result not shown), getting the same result.

In the same conditions (PHD not degraded), we have numerically tested the irreversibility property, but without success: the sharp transition occurred at the same critical value by increasing or decreasing the parameter. It seems more realistic to assume a shift than a switch.

2.5.3 Artfactual sequestration effects

The exhibition of a non-generic switch-like behavior in (Kohn et al., 2004) is an illustration of a more general situation, that we now explain. From a modeling point-of-view, it is not equivalent to assume that a species has a constant concentration or that it is produced and degraded at constant rates. Let us illustrate this phenomenon on VHL. Assume that, to the contrary of our model, VHL is simply a constant parameter. Then, the available concentration of VHL corresponds to VHL which

is not being used in the system (“sequestered”). Since the rate of degradation of $\text{HIF}(\text{OH}_p)_2$ and $\text{HIFOH}_a(\text{OH}_p)_2$ by the proteasome is proportional to the available concentration of VHL, $\text{HIF}(\text{OH}_p)_2$ at equilibrium has a nonlinear dependance on HIF, and $\text{HIFOH}_a(\text{OH}_p)_2$ at equilibrium has a nonlinear dependance on HIFOH_a . This fact has a consequence on HIF_{TOT} , because the dependence on HIFOH_a becomes explicit. In particular, we predict that HIF_{TOT} depends on FIH in this context. By assuming that VHL has non-zero degradation and production rates, the problem disappears since the sum HIF_{TOT} is not anymore dependent on FIH.

Conversely, we can predict that under pathological conditions where production or degradation of proteins (like PHD, FIH or VHL in our model) are stopped or even severely slowed down, sequestration effects can occur and induce nonlinearities in the system, giving unexpected effects like the dependence of HIF_{TOT} on FIH when VHL is affected. The theoretical model would be worth exploiting for deeper studies of this kind of unusual situations. Finally, notice that, in another context, sequestration effects have been proposed as an explanation to unexpected damped temporal oscillations under constant stimulation of signaling cascades (see (Ventura et al., 2008), and also (Legewie et al., 2007)).

3 Discussion

3.1 A theoretical model that is dedicated to biology

We designed a Molecular Interaction Map including key actors of the HIF pathway and major regulators of the gene induction profile downstream of HIF. We attempted to adhere as closely as possible to biological and molecular reality by systematically evaluating our parameters with reference to biological experiments.

From the modeling point-of-view this system has many virtues. First, it reflects all biological features considered in the problem, but stays relatively simple since it admits an analytical treatment which -1- helps in identifying pertinent parameters, $([\text{HIF}]_0, \tau_1, \tau_2, \tau_3)$, and -2- helps in fitting with experiments without requiring random exploration of parameter space. Finally, it is a prediction tool, in the sense that it can be used to quantify the coefficient of sensitivity q of a given gene to the C-TAD domain of HIF, while this measure is not experimentally easily accessible.

From the biological point-of-view, HIF is considered as a key transcription factor driving multiple fundamental biological pathways. In the context of microenvironmental stress adaptation, HIF-targeted genes play crucial functions such as angiogenesis, pH control, extracellular matrix modulation, inflammation, migration or metabolic change (for review (Pouyssegur et al., 2006)). This work introduces an original function of the oxygen sensor FIH (supporting the fact that it is not redundant with the other oxygen sensor PHD). This O_2 -sensitive enzyme allows physical segregation of two potential HIF spectrums in an oxygen gradient: N-TAD-dominant genes would be associated with mild hypoxia (in close vicinity to blood vessels) whereas C-TAD responsive genes would release their maximal expression in strong hypoxia (in periphery of blood vessels). Thus our updated hypoxic MIM indicates how the HIF repertoire can be spatially driven thanks to FIH, and what are the amplitude and the limit of this oxygen sensor.

3.2 Parameter space

The formula for HIF_{TOT} (Equation [3]) allows us to understand what are the parameters which control HIF_{TOT} in an oxygen gradient, namely $[\text{HIF}]_0, a_1, d_1, b_1, [\text{PHD}]_0, l_0$ and τ_3 . More interestingly, the signal HIF_{TOT} actually depends exclusively on $[\text{HIF}]_0$ and the ratios $d_1/b_1, a_1[\text{PHD}]_0/\tau_3$ and τ_3/l_0 . This shows that many combinations of the initial kinetic parameters lead to the same global dynamics in this interaction network, with the same triplet a, b and c . Our interpretation is that this system is potentially able to adapt to punctual kinetic changes and to recover its functionality. For example, if a pharmacological drug inhibits the association between HIF and PHD (decrease in a_1), the system can readapt exactly to the same dynamics by increasing the production of PHD (by a gene duplication for instance), or by counterbalancing τ_3 (enzymatic activity of VHL, that can be altered by single mutations) and l_0 (non specific degradation of HIF). On the other side, perturbations in this dynamical system can also naturally occur in the Von Hippel Lindau disease, generating hematological disorders and kidney tumors. In this context our model should allow evaluation of the impact of different VHL mutations leading to a decreased a_3 , and to show how to restore a physiological context thanks to an adjusted therapy.

3.3 Conclusion

This work illustrates how mathematical tools and biological experiments can synergize in order to get a better understanding of complex interaction networks. In the case of HIF, our data allow us to finely characterize the behavior of this master signaling pathway. Deregulation in this mechanism leads to pathological contexts such as the von Hippel Lindau disease. Our model is a tool that should allow a better understanding of this type of disorder, and also to find ways to therapeutically restore a physiological balance. Our model predicts how FIH allows induction of a precise gene at a precise location within

an oxygen gradient. From a physiological point-of-view, it is highly probable that each gene has to be accurately located in order to properly adapt cells to their microenvironmental stress. This precise spatial regulation should be particularly pertinent within a solid tumour where the dynamics of angiogenesis and environmental adaptation is directly linked to cancer progression. Consequently, as a gene induction organizer, FIH could be an interesting anti-cancer target. It opens the way to an original type of micro-environment directed therapy aiming at disorganizing the spatial profile of the HIF repertoire. In this context, such a type of therapeutical strategy could be implemented, evaluated and optimized in our biomathematical model.

4 Materials and methods

4.1 Approach

We propose to model the regulation of HIF interaction network taking into account both PHD and FIH. The network is represented by a Molecular Interaction Map (MIM, see Figure 2), that is a diagram, with a well defined symbolic vocabulary, which represents all biochemical interactions in a given system. Dynamics are modeled using differential equations (see Figure 11). The system is relatively simple, since we are able to derive the exact values of the equilibrium state in terms of the parameters. This allows a detailed analysis of the evolution of this equilibrium in an oxygen gradient from 21 % to 0 % O_2 .

An essential feature of our model that differs from a previous model in (Kohn et al., 2004), is the lack of switch-like behavior for the HIF concentration as a function of oxygen. Namely, there is no region in parameter space where more than one stable equilibrium state can exist. Moreover we show experimental biological evidence confirming the progressive, without sharp transition behavior of the HIF protein induction.

We explain that the sharp transition previously observed in (Kohn et al., 2004) is an indirect consequence of protein sequestration effect occurring because the system is put in very particular and non-physiological conditions (production and degradation rate of PHD equal to zero).

We illustrate this result by making a simulation within the setting of (Kohn et al., 2004) (see Figure 10).

On the other hand, the analytic expression of the total HIF protein concentration (Equation in Lemma 2.1) allows to exhibit three new parameters which are combinations of the initial ones, which entirely determine quantitatively the HIF concentration as a function of O_2 . Using experimental data we fit the model to the experiment. This model gives new insight into the way FIH can modulate the spectrum of HIF-targeted adaptation genes. In the context of a virtual oxygen gradient, it clearly shows that FIH is able to displace the C-TAD repertoire into highly hypoxic areas while N-TAD genes express in mildly hypoxic zones. In accordance with biological reality, this mechanism leads to a better understanding of how the oxygen sensors drive the gene expression pattern from a blood vessel to its hypoxic periphery.

Notice that three isoforms of the α subunit have been characterized. Here we consider the presence of only one isoform that we call HIF α . Note that the cell line that we use (LS174) is indeed a HIF-1 α only cell line (data not shown).

4.2 Building the models

4.2.1 Molecular Interaction Map

Basing on a previous MIM (Kohn et al., 2004), the originality of our model relies in three aspects:

- Based on the fact that HIF bears two distinct and functional transactivation domains (TAD) (Dayan et al., 2006), (Gothie et al., 2000), our hypothesis is that both the N-TAD and the C-TAD can potentially stimulate the spectrum of HIF downstream genes. In order to differentiate these two domains, we chose to assess the sensitivity to the C-TAD with a sensitivity coefficient $q \geq 0$, which is characteristic of a given gene, with a given HRE. When FIH hydroxylates HIF on its Asn-residue target, this C-TAD related stimulation is blocked. A hypothetical HRE with $q = 0$ does not respond to the C-TAD domain. Its regulation is not sensitive to FIH, and entirely controlled by PHD. On the other side, a HRE with $q > 0$ is activated by both transactivation domains of HIF: its regulation is strongly controlled by FIH, and q reflects the amplitude of this modulation.
- In agreement with the molecular description of hydroxylation of HIF by PHD, we take into account that HIF is hydroxylated on two proline residues and the reaction involves two dioxygen molecules (Masson et al., 2001), (Chan et al., 2005).
- we assume that the dynamics of the β subunit of HIF (species ARNT in (Kohn et al., 2004)) is not relevant in itself, and is not limiting for the other reactions.

In Table 1, we list the variables of our model. Figure 2 is the underlying MIM. The parameters are:

- Kinetic parameters: association/dissociation constants a_i, d_i , conversion constants (k_i, b_i) , degradation constants (l_i) . In Table 2 C), we give a set of parameter values obtained by fitting the model to experimental data (see Sections 2.1.3 and 2.4). See also Table 2 B) for units;
- Molecular species: Precursors of HIF (S_0), PHD (S_1), FIH (S_2), VHL (S_3); Oxygen ($[O_2]$).

In the following, we always suppose that, otherwise stated, all parameters are non zero.

4.2.2 Dynamical equations

From the static picture given by the MIM, we derive the system of differential equations which models the dynamics. Each reaction in the MIM is supposed to follow the principle of mass action: its rate is proportional to the product of the concentrations of the reactants to the power of the stoichiometric coefficient. Equations are obtained by making a mass balance for each vertex of the MIM. The equations are given in Figure 11.

Remark 4.1 Since we have assumed the double hydroxylation of HIF by PHD on Pro 402 and Pro 564 (see 2.1.4 for further analysis of prolyl hydroxylation, and (Schofield and Ratcliffe, 2004)), O_2 appears as a second order in the rates of the elementary reactions involved in this hydroxylation. For that reason, one encounters the term $[O_2]^2$ in the equations.

Remark 4.2 Species HRE plays a particular role, due to the fact that there is usually one such element per cell and it cannot be considered to evolve in time as a continuous variable. To avoid this difficulty, it is typical to assume that we are modeling the mean behavior of a population of identical cells. Then, the variable x_4 stands for the proportion of HIF-free HRE in the population, x_{13} , x_{15} , x_{14} and x_{16} are proportions of HRE bounded to HIF, HIFOH_a, HIF(OH_p)₂ and HIFOH_a(OH_p)₂ respectively. They are real numbers between 0 and 1, and their sum is 1 (because we suppose that the total amount of HRE is constant).

4.2.3 Numerical fitting

We evaluated the numerical values of the parameters from experimental data by nonlinear fitting with the help of the software *Mathematica*® [Wolfram Research, Inc., Mathematica, Version 6, Champaign, IL (2007)].

4.3 Quasi-steady state model reduction

The quasi-steady state model reduction consists in applying the standard Briggs-Haldane quasi-steady state approximation (see, for instance (Briggs and Haldane, 1925)) For instance, consider the hydroxylation of HIF by PHD: our initial model decomposes this process into two reactions, namely the formation/dissociation of the complex HIF:PHD with rates a_1 and d_1 respectively, and the binding with $[O_2]$ with release of PHD, with rate b_1 . If one assumes that $d[\text{HIF:PHD}]/dt = 0$ (Briggs-Haldane approximation), then hydroxylation is well represented by one reaction, with a rate of production, say V_0 , of HIF(OH_p)₂ which follows a Michaelis-Menten equation:

$$V_0 = \frac{V_{\max}^1 [\text{HIF}]}{K_M^1 + [\text{HIF}]}$$

and we have:

$$V_{\max}^1 = b_1 [\text{PHD}] [O_2]^2, \quad K_M^1 = \frac{d_1 + b_1 [O_2]^2}{a_1}$$

Let us define the following parameter:

$$\tau_1 = \frac{V_{\max}^1}{K_M^1}$$

which is the steepness of the Michealis-Menten curve near the origin. This parameter appears naturally in the sequel.

In the same way, we define the parameters τ_2 , which is related to the hydroxylation of HIF by FIH, and τ_3 , related to the degradation through VHL by the following formulae:

$$V_{\max}^2 = b_2 [\text{FIH}] [O_2], \quad K_M^2 = \frac{d_2 + b_2 [O_2]}{a_2}, \quad \tau_2 = \frac{V_{\max}^2}{K_M^2}$$

and

$$V_{\max}^3 = b_3 [\text{VHL}], \quad K_M^3 = \frac{d_3 + b_3}{a_3}, \quad \tau_3 = \frac{V_{\max}^3}{K_M^3}$$

Figure 3 illustrates diagrammatically the quasi-steady state model reduction.

Remark 4.3 Recall that in enzymology (see for instance (Cornish-Bowden, 1995)) the *specificity constant* measures the efficiency of an enzyme in converting a substrate into product. For the enzyme PHD and substrate HIF it equals:

$$k_{\text{HIF}}^{\text{PHD}} = \frac{b_1[\text{O}_2]^2}{K_M^1} = \frac{a_1 b_1 [\text{O}_2]^2}{d_1 + b_1 [\text{O}_2]^2}$$

In this context

$$\tau_1 = k_{\text{HIF}}^{\text{PHD}} [\text{PHD}].$$

Analogously, we can introduce the specificity constant of FIH for HIF and the one of VHL for HIF(OH_p)₂ and HIFOH_a(OH_p)₂:

$$k_{\text{HIF}}^{\text{FIH}} = \frac{b_2[\text{O}_2]}{K_M^2} = \frac{a_2 b_2 [\text{O}_2]}{d_2 + b_2 [\text{O}_2]}$$

$$k_{\text{HIF(OH}_p)_2}^{\text{VHL}} = \frac{b_3}{K_M^3} = \frac{a_3 b_3}{d_3 + b_3}$$

so that the coefficients τ_2 and τ_3 read:

$$\tau_2 = k_{\text{HIF}}^{\text{FIH}} [\text{FIH}] \quad \tau_3 = k_{\text{HIF(OH}_p)_2}^{\text{VHL}} [\text{VHL}]$$

4.4 Experimental biology

4.4.1 Immunoblotting and quantification

LS174 cells were lysed in 1.5x Laemmli buffer. The protein concentration was determined using the bicinchoninic acid assay. Whole-cell extracts (40 μ g for HIF-1 α) were resolved by SDS-PAGE (7.5%) and transferred onto a polyvinylidene difluoride membrane (Millipore, Bedford, MA). Immunoreactive bands were visualized with the enhanced chemiluminescence system (Amersham Biosciences, Buckinghamshire, UK).

Anti-HIF-1 α (antiserum 2087) and anti-Erk (antiserum E1B4) were produced and characterized previously (Richard et al., 1999), (Brondello et al., 1999).

The intensity of the specific HIF-1 α band was quantified using the GeneTool program (Syngene / Synoptics, Cambridge, UK). The loading control was done with the anti-Erk antibody.

4.4.2 Cell culture and cell line production

Genetically modified cell lines, overexpressing or invalidating FIH (shRNA) were previously described in (Dayan et al., 2006).

4.4.3 RNA Extraction

Total RNA was extracted from LS174 cells using Trizol reagent (Life Technologies) according to the manufacturer's instructions. Briefly, cultured cells were homogenized in 1 mL Trizol, centrifuged after adding chloroform and precipitating RNA with isopropanol, and washed with 75% ethanol. The RNA purity was evaluated by spectrophotometry. Total RNA (2 μ g) from LS174 cells was added to a 20 μ L reverse transcription-PCR (RT-PCR) reaction using the Omniscript kit (Qiagen, Inc., Valencia, CA).

4.4.4 Quantification of gene expression by real-time PCR

The relative expression level of the series of HIF-downstream targets was quantified by real-time RT-PCR using the Taqman PCR Master Mix (Eurogentec) on an ABI Prism 7300 Sequence Detection System (Applied Biosystems, Foster City, CA) according to the manufacturer's instructions. For each gene, the relative induction (with 36B4 as a reference gene) was calculated by using the equation $2^{-\Delta\Delta\text{CT}}$, where $\Delta\text{CT} = \text{CT}(\text{gene}) - \text{CT}(36\text{B4})$ and $\Delta\Delta\text{CT} = \Delta\text{CT}(\text{stimulated condition}) - \Delta\text{CT}(\text{unstimulated condition})$. Each gene was amplified using the appropriate specific primers. Each individual induction is the mean of two amplifications.

Appendix

Dynamical model

The differential equations governing the MIM in Figure 2 are given in the following Figure 11.

Equilibrium state analysis

Equilibria of the system are solutions to the set of equations

$$\frac{dx_i}{dt} = 0 \quad i = 0, \dots, 16.$$

We suppose that all parameters are non-zero. In the following, we show that the equilibrium equations can be solved analytically: the equilibrium values can be expressed as functions of the parameters of the system. We denote e_i the equilibrium value of the variable x_i .

First, by summing Eq. (1), (8) and (11) in Fig. 11, we find the equilibrium value for [PHD]:

$$e_1 = [\text{PHD}]_0 = \frac{k_1 S_1}{l_1} \quad (9)$$

Then, the equilibrium value for [FIH] is obtained by summing Eq. (2) and (9) in Fig. 11:

$$e_2 = [\text{FIH}]_0 = \frac{k_2 S_2}{l_2} \quad (10)$$

To find the equilibrium value of [VHL], we sum Eq. (3), (10) and (12) in Fig. 11:

$$e_3 = [\text{VHL}]_0 = \frac{k_3 S_3}{l_3} \quad (11)$$

We express the equilibrium values of the intermediary products in terms of parameters and the other variables by solving $\dot{x}_8 = \dot{x}_9 = \dot{x}_{10} = \dot{x}_{11} = \dot{x}_{12} = 0$.

$$\begin{aligned} e_8 &= \frac{a_1 e_5 e_0}{d_1 + b_1 [\text{O}_2]^2} & e_9 &= \frac{a_2 e_6 e_0}{d_2 + b_2 [\text{O}_2]} & e_{10} &= \frac{a_3 e_{16} e_3}{d_3 + b_3} \\ e_{11} &= \frac{a_1 e_5 e_4}{d_1 + b_1 [\text{O}_2]^2} & e_{12} &= \frac{a_3 e_{16} e_2}{d_3 + b_3} \end{aligned}$$

Using equations [1], [5], [2] for the definition of τ_1 , τ_2 and τ_3 we get

HIF	$e_0 = \frac{k_0 S_0}{l_0 + \tau_1 ([\text{O}_2]) + \tau_2 ([\text{O}_2])}$
HIF(OH _p) ₂	$e_5 = \frac{\tau_1}{\tau_3} e_0$
HIFOH _a	$e_6 = \frac{\tau_2}{l_0 + \tau_1} e_0$
HIFOH _a (OH _p) ₂	$e_7 = \frac{\tau_1}{\tau_3} \frac{\tau_2}{l_0 + \tau_1} e_0$

The function $[\text{HIF}_{\text{TOT}}] = e_0 + e_5 + e_6 + e_7$ reads:

$$[\text{HIF}_{\text{TOT}}] = \frac{k_0 S_0}{l_0} \frac{1 + \frac{\tau_1 ([\text{O}_2])}{\tau_3}}{1 + \frac{\tau_1 ([\text{O}_2])}{l_0}}$$

Remark A.1 If one of the degradation rates l_1 or l_2 is 0, for instance $l_1 = 0$, then Equation [9] is no more valid. There are two possibilities:

- If $k_1 S_1$ is not zero, then PHD keeps on accumulating, and does not reach any equilibrium.
- If $k_1 S_1 = 0$, then, for all time, $\frac{dx_1}{dt} + \frac{dx_8}{dt} + \frac{dx_{11}}{dt} = 0$ that is the equilibrium value of PHD depends on HIF and Equation [3] in Lemma 2.1 is not valid anymore.

The equilibrium equations for the HRE and its complexed forms with HIF read:

HIF:HRE	$e_{13} = \frac{a_4}{d_4} e_4 e_0$
HIF(OH _p) ₂ :HRE	$e_{14} = \frac{a_4}{d_4} e_5 e_4$
HIFOH _a :HRE	$e_{15} = \frac{a_4}{d_4} e_6 e_4$
HIFOH _a (OH _p) ₂ :HRE	$e_{16} = \frac{a_4}{d_4} e_7 e_4$

The quantities e_0, e_5, e_6 and e_7 have been computed before. All HRE complexes have been expressed as functions of free HRE, that is e_4 . In order to compute e_4 , we use the relation $e_4 = 1 - e_{13} - e_{14} - e_{15} - e_{16}$ (See Remark 4.2). It gives:

$$e_4 = \frac{1}{1 + \frac{a_4}{d_4} (e_0 + e_5 + e_6 + e_7)}$$

For a q -gene, the stimulation function $G_q = e_{13} + e_{14} + e_{15} + e_{16} + q(e_{13} + e_{14})$ (see Equation [4]) equals:

$$G_q([O_2]) = \frac{a_4}{d_4} e_4 ([HIF_{TOT}]([O_2]) + q(e_0 + e_5))$$

It can be written in terms of the coefficients τ_i (Equation [6]):

$$G_q([O_2]) = \frac{a_4}{d_4} \frac{k_0 S_0}{l_0} \frac{1 + \frac{\tau_1}{\tau_3}}{1 + \frac{\tau_1}{l_0} + \frac{a_4}{d_4} \frac{k_0 S_0}{l_0} (1 + \frac{\tau_1}{\tau_3})} (1 + q \frac{1 + \frac{\tau_1}{\tau_3}}{1 + \frac{\tau_1}{l_0} + \frac{\tau_2}{l_0}})$$

Fitting the stimulation function and coefficient q

We use the coefficients a , b and c estimated in the fitting of the function h .

Expressing g_q (Equation [7]) in terms of those coefficients gives:

$$g_q = \frac{100c(1 + \frac{a_4}{d_4} [HIF]_0)([O_2]^2 + b)}{(q + 1)((b + c \frac{a_4}{d_4} [HIF]_0)[O_2]^2 + bc(1 + \frac{a_4}{d_4} [HIF]_0))} \cdot (1 + q \frac{\delta_1([O_2]^2 + c)([O_2] + \frac{d_2}{b_2})}{\delta_1([O_2]^2 + c)([O_2] + \frac{d_2}{b_2}) + c([O_2]^2 + \delta_1) \frac{a_2 [FIH]_0}{l_0} [O_2]})$$

which is Equation [8].

When the ratio of induction at some $p[O_2]$ is known, it is possible to extract the expression of q from the equation:

$$g_q(x) = y$$

where y is the ratio of induction at $x = 3\%$. We get an expression of the form:

$$q = \frac{y_0 - y}{y - y_1}$$

where y_0 is the induction rate for a N-TAD-only gene:

$$y_0 = \frac{(100(9 + b)c(1 + \rho))}{(9c\rho + b(9 + c + c\rho))}$$

and:

$$y_1 = \frac{80(9 + c)\delta_1(9 + \delta_2)}{27c\epsilon_1 + \delta_1(9(3 + \delta_2) + c(3 + \delta_2 + 3\epsilon_1))}$$

The quantity y_1 represents the vertical asymptote value of the graph of q .

We identify the following set of independent parameters which remain to be fitted:

$$q, \quad \rho = \frac{a_4}{d_4} [HIF]_0, \quad \epsilon_1 = \frac{a_2 [FIH]_0}{l_0}, \quad \delta_1 = \frac{d_1}{b_1}, \quad \delta_2 = \frac{d_2}{b_2}$$

We choose the maximum induction rate $x_0 = 80\%$ and the minimum induction rate $x_4 = 3.5\%$, from which we derive:

$$\rho = \frac{36b - 45c - bc}{(9 + b)c}$$

and

$$\epsilon_1 = \frac{\delta_1(4257 + 473c + 459\delta_2 + 51c\delta_2)}{c(63 + 7\delta_1)}$$

Experimental induction of the gene *ca-IX* give three data points, which we use to numerically fit δ_1 and δ_2 and q_{ca-IX} . The result of Mathematica[©] fitting is:

$$\delta_1 = 15269.2, \quad \delta_2 = 6498.53, \quad q_{CA9} = 3.60055$$

We admit that ρ , ϵ_1 and δ_2 are independent on q , that is on the specific gene used for the fit, and use the preceding values and the experimental data from quantitative PCR performed on a series of 25 genes to evaluate their C-TAD-sensitivity coefficient q .

Aknowledgements

We thank both referees of the first version of the paper for their helpful and detailed comments.

References

- Arkin, A., Ross, J., and Mac Adams, H. (1998). Stochastic kinetic analysis of developmental pathway bifurcation in phage λ -infected escherichia coli cells. *Genetics*, 149(4):1633–1648.
- Brahimi-Horn, M. and Pouyssegur, J. (2007). Harnessing the hypoxia-inducible factor in cancer and ischemic disease. *Biochem Pharmacol*, 73(3):450–457.
- Briggs, G. E. and Haldane, J. B. S. (1925). A note on the kinetics of enzyme action. *Biochemical Journal*, 19(2):338–339.
- Brondello, J., Pouyssegur, J., and McKenzie, F. (1999). Reduced MAP kinase phosphatase-1 degradation after p42/p44 MAPK-dependent phosphorylation. *Science*, 286(5449):2514–2517.
- Chan, D., Sutphin, P., and Yen, S.E. and Giaccia, A. (2005). Coordinate regulation of the oxygen-dependent degradation domains of hypoxia-inducible factor 1 α . *Mol Cell Biol*, 25(15):6415–6426.
- Cornish-Bowden, A. (1995). *Fundamentals of enzyme kinetics*. Portland Pr.
- Dayan, F., Mazure, N., Brahimi-Horn, M. C., and Pouyssegur, J. (2008). A dialogue between the hypoxia-inducible factor and the tumor microenvironment. *Cancer Microenvironment*, 1(1):53–68.
- Dayan, F., Roux, D., Brahimi-Horn, M. C., Pouyssegur, J., and Mazure, N. (2006). The oxygen sensor factor-inhibiting hypoxia-inducible factor-1 controls expression of distinct genes through the bifunctional transcriptional character of hypoxia-inducible factor-1 α . *Cancer Res*, 66(7):3688–3698.
- Folkman, J. (2007). Angiogenesis: an organizing principle for drug discovery? *Nat Rev Drug Discov*, 6(4):273–286.
- Gothie, E., Richard, D., Berra, E., Pagès, G., and Pouyssegur, J. (2000). Identification of alternative spliced variants of human hypoxia-inducible factor-1 α . *J Biol Chem*, 275(10):6922–6827.
- Ivan, M., Kondo, K., Yang, H., Kim, W., Valiando, J., Ohh, M., Salic, A., Asara, J., Lane, W., and Kaelin, W. (2001). Hif α targeted for vhl-mediated destruction by proline hydroxylation: implications for o₂ sensing. *Science*, 292(5516):464–468.
- Jaakkola, P., Mole, D., Tian, Y., Wilson, M., Gielbert, J., Gaskell, S., Kriegsheim, A., Hebestreit, H., Mukherji, M., Schofield, C., Maxwell, P., Pugh, C., and Ratcliffe, P. (2001). Targeting of hif α to the von hippel-lindau ubiquitylation complex by o₂-regulated prolyl hydroxylation. *Science*, 292(5516):468–472.
- Kohn, K., Riss, J., Aprelikova, O., Weinstein, J., Pommier, Y., and Barrett, J. (2004). Properties of switch-like bioregulatory networks studied by simulation of the hypoxia response control system. *Mol Biol Cell*, 15(7):3042–3052.
- Lando, D., Peet, D., Gorman, J., Whelan, D., Whitelaw, M., and Bruick, R. (2002a). Fih-1 is an asparaginyl hydroxylase enzyme that regulates the transcriptional activity of hypoxia-inducible factor. *Genes Dev*, 16(12):1466–1471.
- Lando, D., Peet, D., Whelan, D., Gorman, J., and Whitelaw, M. (2002b). Asparagine hydroxylation of the hif transactivation domain a hypoxic switch. *Science*, 295(5556):858–861.
- Legewie, S., Schoeberl, B., Blthgen, N., and Herzel, H. (2007). Competing docking interactions can bring about bistability in the mapk cascade. *Biophys J*, 93(7):2279–2288.
- Mahon, P., Hirota, K., and Semenza, G. (2001). Fih-1: a novel protein that interacts with hif-1 α and vhl to mediate repression of hif-1 transcriptional activity. *Genes Dev*, 15(20):2675–2686.
- Masson, N., Willam, C., Maxwell, P., Pugh, C., and Ratcliffe, P. (2001). Independent function of two destruction domains in hypoxia-inducible factor- α chains activated by prolyl hydroxylation. *EMBO J*, 20(18):5197–5206.
- Monod, J. and Jacob, F. (1961). General conclusions: teleonomic mechanisms in cellular metabolism, growth, and differentiation. In *Cold Spring Harbor Symp Quant Biol*, volume 26, pages 389–401.

- Pouyssegur, J., Dayan, F., and Mazure, N. (2006). Hypoxia signalling in cancer and approaches to enforce tumour regression. *Nature*, 441(7092):437–443.
- Qutub, A. and Popel, A. (2006). A computational model of intracellular oxygen sensing by hypoxia-inducible factor hif-1alpha. *J Cell Sci*, 119(16):3467–3480.
- Richard, D., Berra, E., and Gothie, E. (1999). p42/p44 mitogen-activated protein kinases phosphorylate hypoxia-inducible factor 1a (hif-1a) and enhance the transcriptional activity of HIF-1. *J Biol Chem*, 274(46):32631–32637.
- Schofield, C. J. and Ratcliffe, P. J. (2004). Oxygen sensing by hif hydroxylases. *Nat Rev Mol Cell Biol*, 5(5):343–354.
- Semenza, G. (2003). Targeting HIF-1 for cancer therapy. *Nat Rev Cancer*, 3(10).
- Ventura, A., Sepulchre, J.-A., and Merajver, S. (2008). A hidden feedback in signaling cascades is revealed. *PLoS Comput Biol*, 4(3):e1000041.
- Yan, Q., Bartz, S., Mao, M., Li, L., and Kaelin, W. J. (2007). The hypoxia-inducible factor 2alpha n-terminal and c-terminal transactivation domains cooperate to promote renal tumorigenesis in vivo. *Moll Cell Biol*, 27(6):2092–2102.

Figure Legends

Figure 1 Structural view of the regulation of HIF by the two oxygen sensors. In the presence of oxygen, HIF α is hydroxylated by two enzymes called oxygen sensors. The hydroxylation on proline by PHD promotes the interaction between HIF and pVHL, leading to massive proteasomal degradation (Jaakkola et al., 2001), (Ivan et al., 2001). Moreover FIH hydroxylates an asparagine residue on the C-TAD that results in an inactivation of this domain (Lando et al., 2002a), (Lando et al., 2002b); (Mahon et al., 2001). In such conditions the N-TAD is the only domain carrying the activity of HIF. In the absence of oxygen PHD and FIH are inhibited and consequently HIF is stabilized and active *via* its two TADs.

Figure 2 Molecular Interaction Map. S_0 stands for HIF precursor. We have added a precursor to VHL and a degradation reaction to avoid non linear effects due to artifactual protein sequestration (See Paragraph 2.5); S_1 is the PHD precursor; S_2 is the FIH precursor.

Figure 3 The quasi-steady state model reduction. All intermediary complexes have been eliminated. Parameters τ_1 , τ_2 and τ_3 are defined by Equations [1], [5] and [2], respectively. See also Section 4.3 for explanation.

Figure 4 [HIF] versus $[O_2]$ variation. (A) The ratio l_0/τ_3 reflects the equilibrium between non specific and oxygen dependent (i.e proteasomal) degradation of HIF. When this ratio is < 1 , [HIF] is a decreasing function of oxygen. This corresponds to the physiological context, when τ_3 (proportional to the oxygen dependent degradation rate of HIF by VHL, see Equation [2]) is greater than the non specific degradation rate l_0 . On the opposite, if these two degradation pathways are equivalent, [HIF] is no longer dependent on $[O_2]$. Finally, when this ratio is > 1 , [HIF] is artificially an increasing function of oxygen. (B) The sharpness of the hypoxic stabilisation of HIF is highly dependent on a_1 , the association constant between HIF and PHD.

Figure 5 Quantification of total HIF protein (A) Experimental quantification of [HIF] in response to hypoxia. Lower panel: the induction of HIF is detected by immunoblotting. LS174 cells (colon carcinoma cell line) were incubated for 24h at different percentages of $[O_2]$ as indicated. Upper panel: based on immunoblotting, the quantification of the HIF α protein induction (using the *Gene Tool(R)* program) is plotted as a function of oxygen. AU = Arbitrary Unit. Note that this quantification is normalized by a hypoxia-independent loading control (not shown). These data are at the basis of our fitting calculation. (B) Nonlinear fitting of $[HIF_{TOT}]$, with data shown on the left, realized with Mathematica[®].

Figure 6 Quantification of the HIF-targeted gene *ca-IX* The nonlinear fitting of the normalized stimulation function g_q (Equation [7]) of the gene *ca-IX* was realized with Mathematica[®]. The score of the fitting reads $d = 6.08$. Red dots represent represent measures from QPCR experiments.

Figure 7 HIF-targeted gene stimulation: FIH relocates C-TAD-sensitive genes in strongly hypoxic areas. Numerical simulations are run using parameter values given in Table 2, that are optimized according to experimental data and thus reflect the physiological cellular context. Gene induction is plotted as a function of oxygen for different q values (i.e. different C-TAD sensitivities). Maximal inductions are normalized to 100. Starting from the arbitrary value of $[FIH] = 100$ we simulated an overexpression and an inhibition of FIH. Dotted lines and arrows indicate the % of oxygen associated to a fixed threshold of gene induction for N-TAD only genes ($q = 0$) and for (N+C)-TAD-sensitive genes (here for $q = 100$). These experiments show how FIH modulates the profile of HIF-targeted genes. Its activity leads to a release of the C-TAD-sensitive genes in highly hypoxic regions while N-TAD dominant genes are localized in mildly hypoxic zones. The level of FIH directly impacts on the amplitude of this gene modulation.

Figure 8 Nonlinear fitting of q , realized with Mathematica[®]. The blue squares correspond to genes previously qualified as FIH-sensitive genes, while green diamonds correspond to non FIH-inhibited genes (see (Dayan et al., 2006)). FIH-sensitive genes have the highest coefficient q . The red dotted line represents the lower limit for the percentage of gene induction at 3% $[O_2]$. See Equation [4] for the exact definition of q , and the last section of Appendix for the numerical fitting of the sensitivity coefficient q . Inset : table of calculated q for each gene.

Figure 9 Confrontation of FIH modulation experiments vs estimated q . The C-TAD sensitivity coefficient q (Equation [4]) is calculated using our model and experimental data coming from partial hypoxia experiments (see Table 1, (Dayan et al., 2006); FIH sensitivity score is calculated on the basis of an independent experiment using genetic modulation of FIH (Dayan et al., 2006). The correlation between these two quantities corroborates the validity of our model, and allows to classify these genes according to their preferential N-TAD (green diamonds, cluster A) or C-TAD (blue squares, clusters B and C) dominancy. Note that *phd3*, *trefoil factor 3*, *p21* and *bnip3* escape this classification and that consequently they are likely to be regulated by extra-factors that are not present in our MIM, or even unknown.

Figure 10 Switch-like behavior is not generic. Simulations are made for $[FIH]_0 = 0$. A switch-like behavior is obtained only for production and degradation rates of PHD equal to 0. Moreover, no irreversibility, and thus no switch, is numerically observed. Parameters and initial concentrations are taken from (Kohn et al., 2004).

Figure 11 Dynamical system. Dynamical equations are built on the basis of the MIM (Figure 2) by applying the mass-action principle to each reaction. The dot over each variable on the left hand side denotes the time derivative of the variable. O_2 is second order in all equations that account for hydroxylation of HIF by PHD, because this hydroxylation is supposed to

occur on two sites: see Remark 4.1 and Section 2.1.4. This biological hypothesis is confirmed by numerical optimization on our data (Table 3). See Table 1 for species' names.

Accepted manuscript

FIGURES

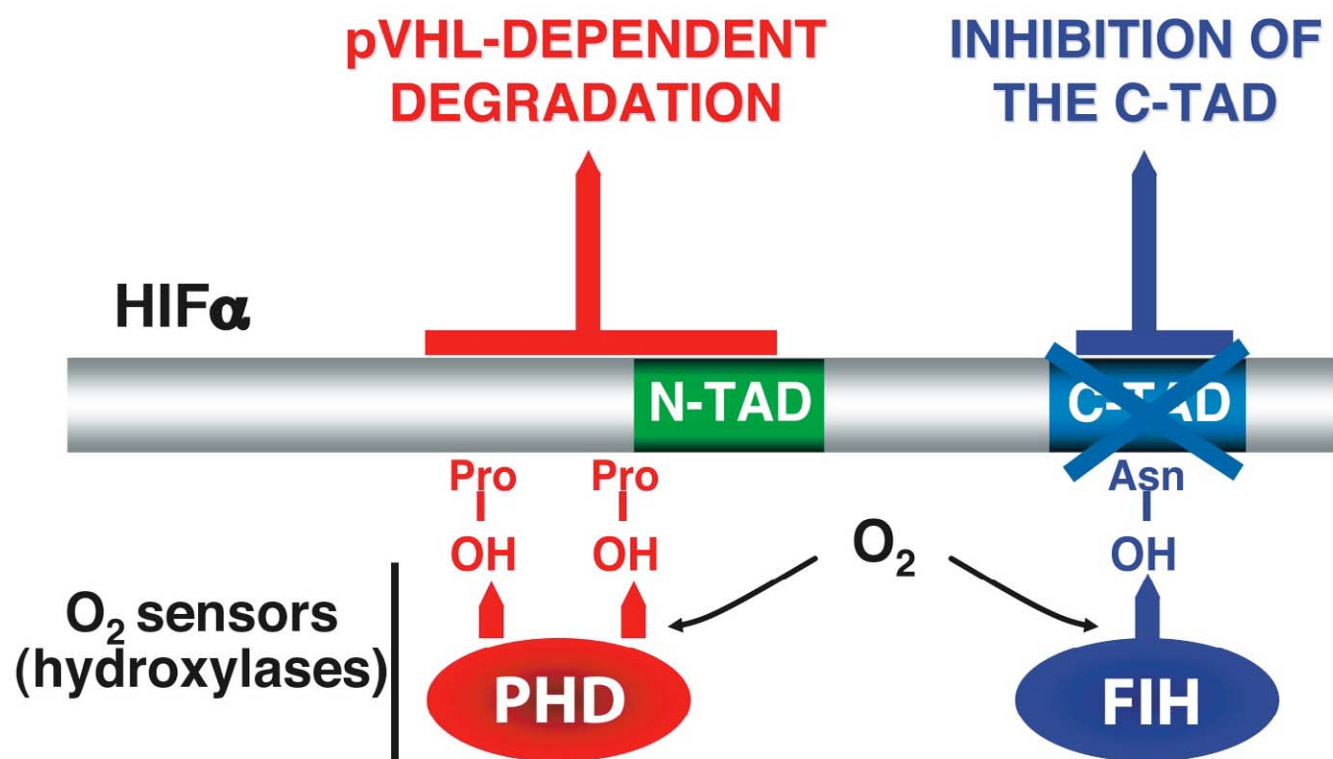


Figure 1:

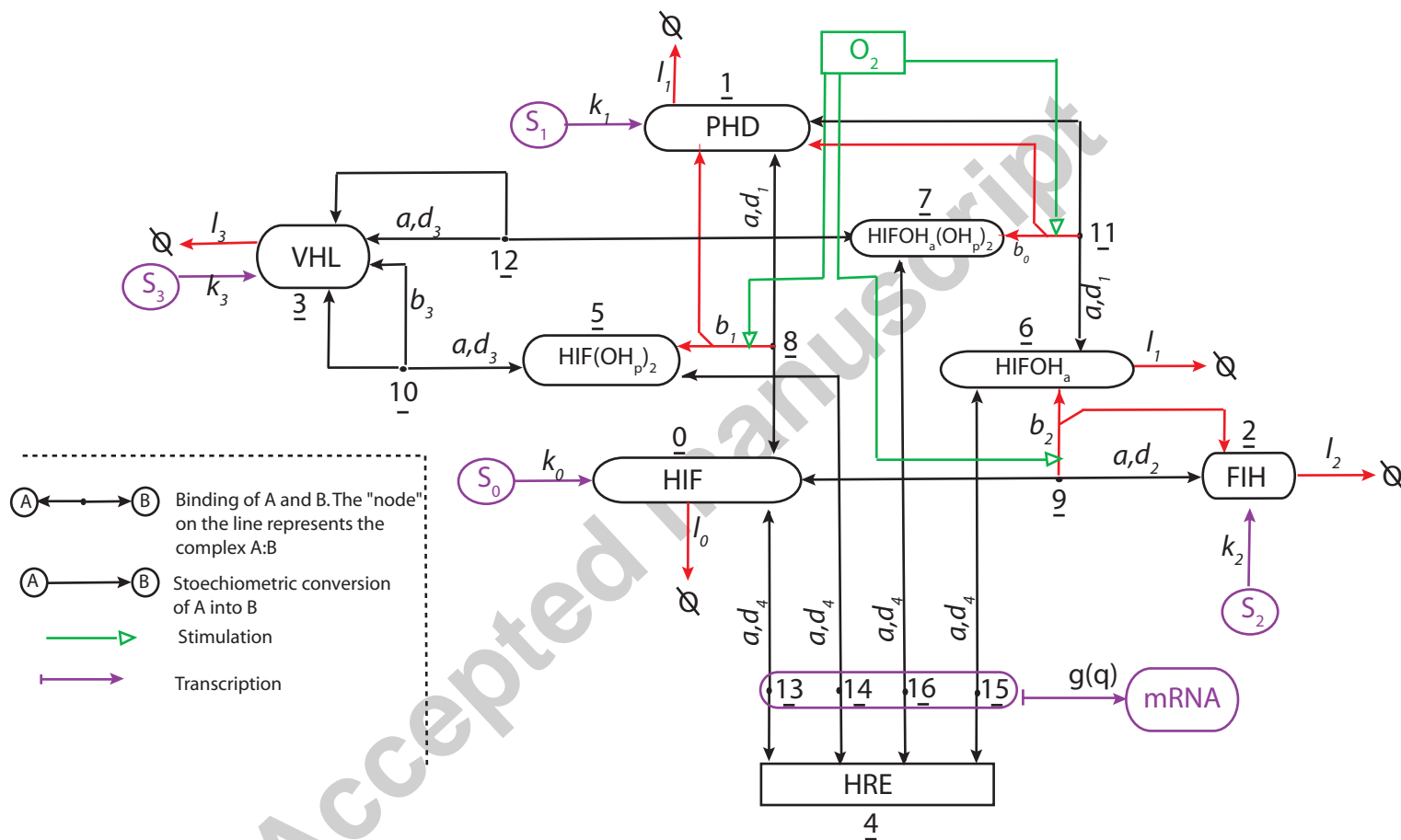


Figure 2:

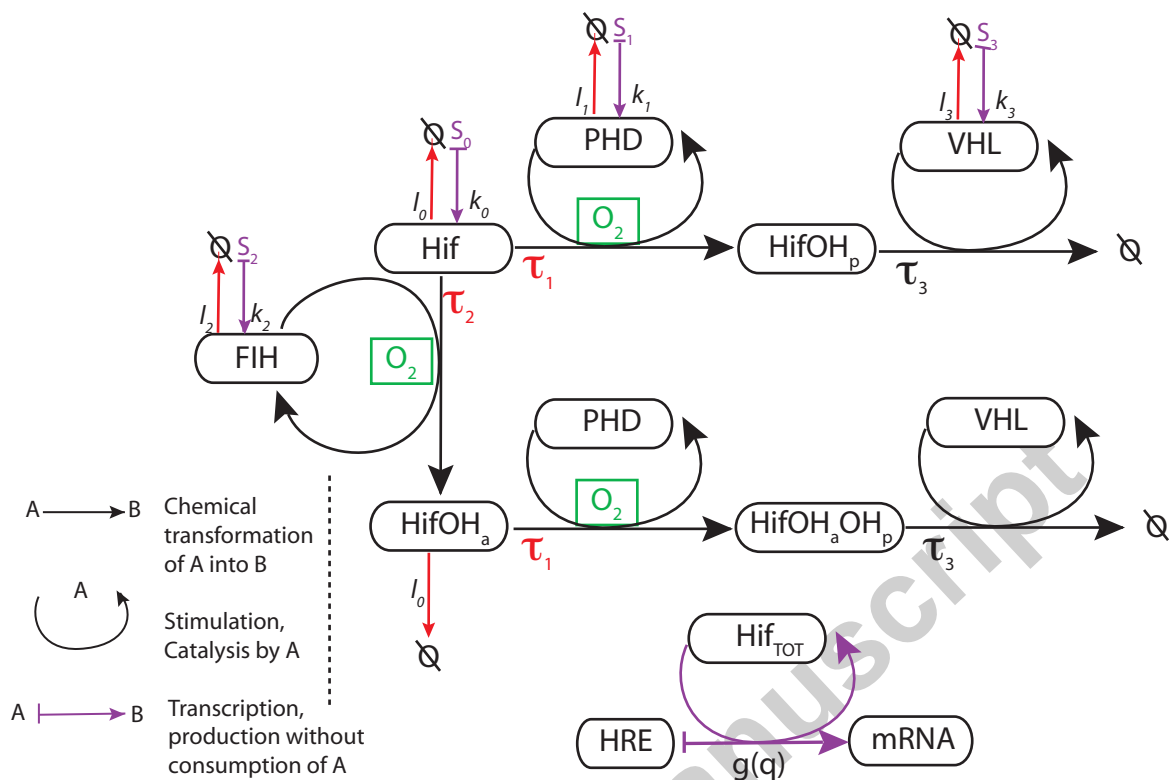


Figure 3:

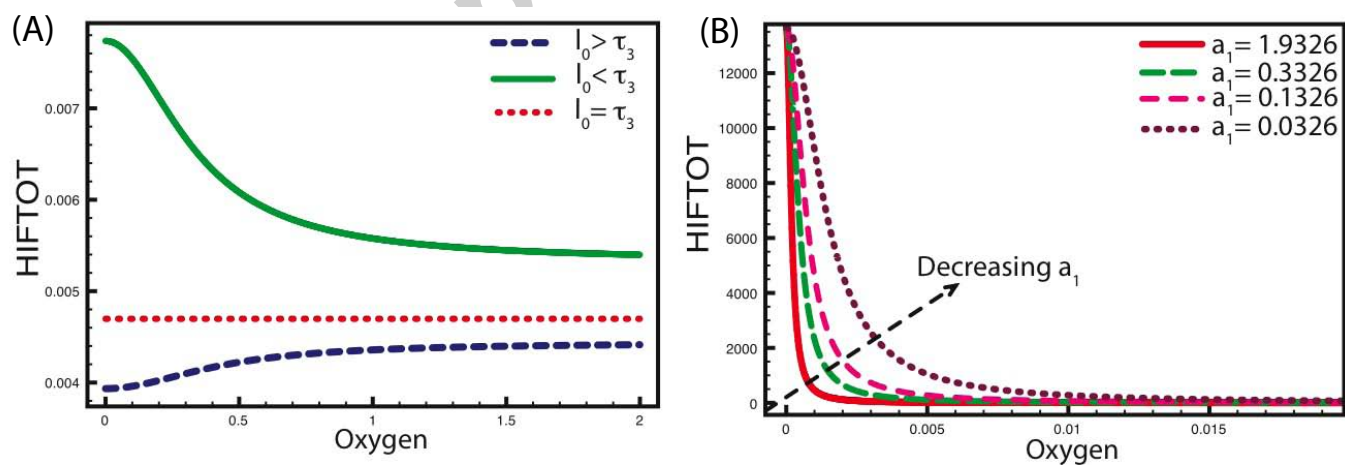


Figure 4:

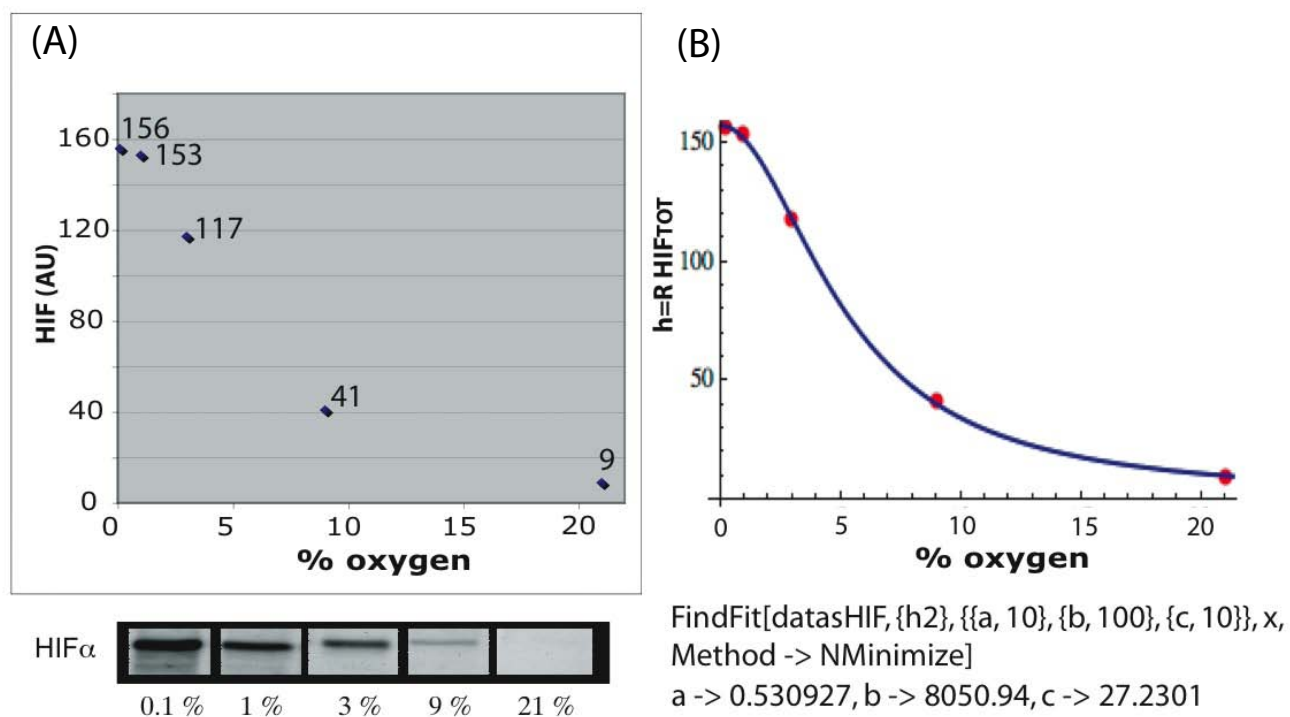


Figure 5:

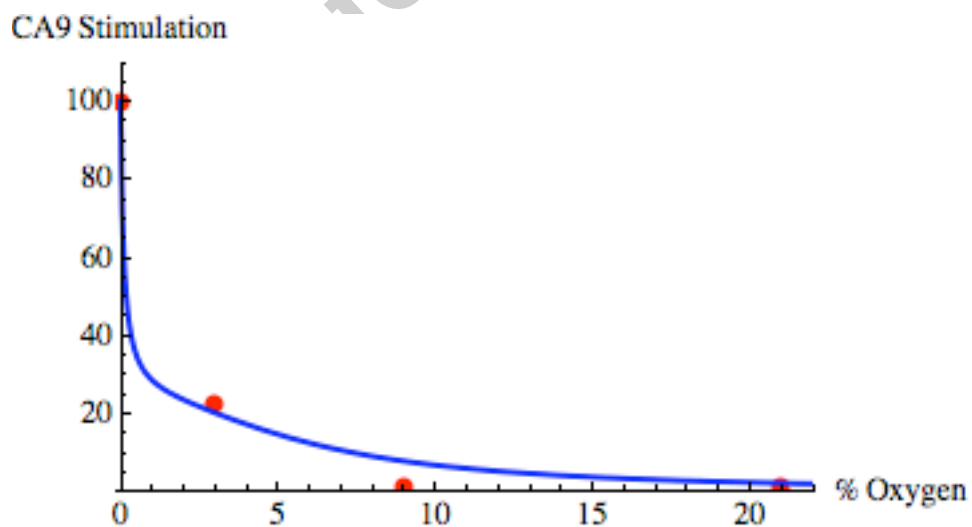


Figure 6:

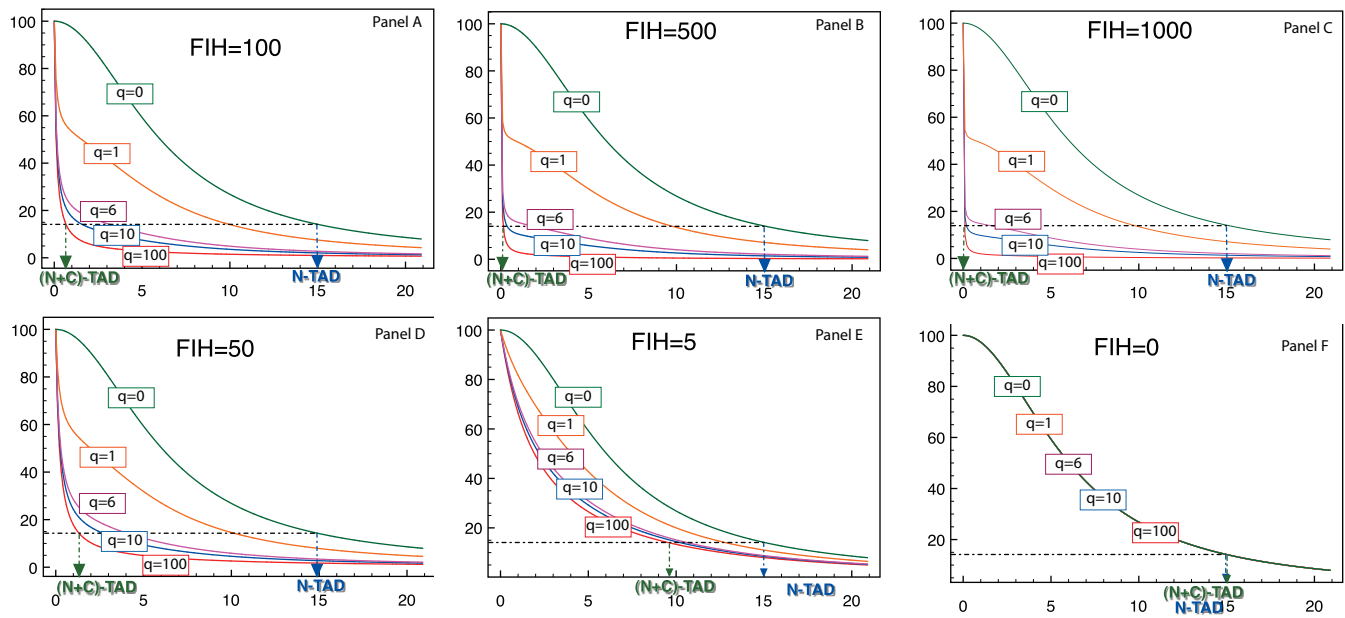


Figure 7:

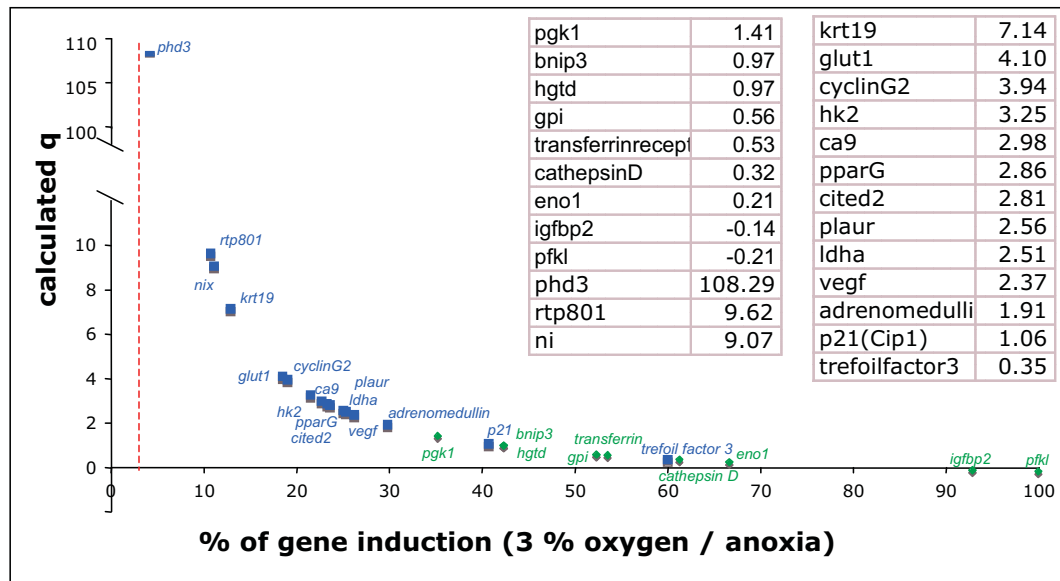


Figure 8:

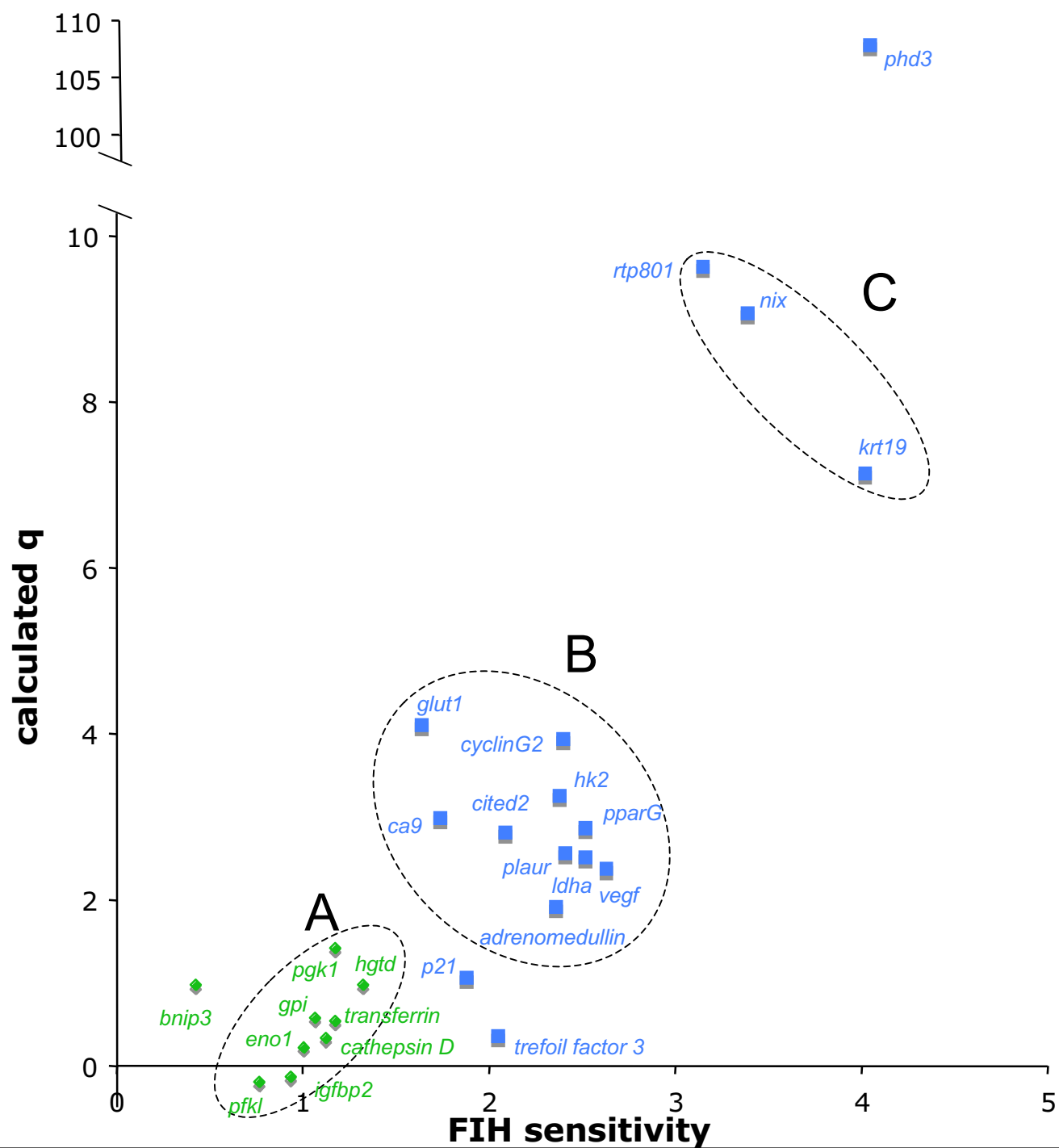


Figure 9:

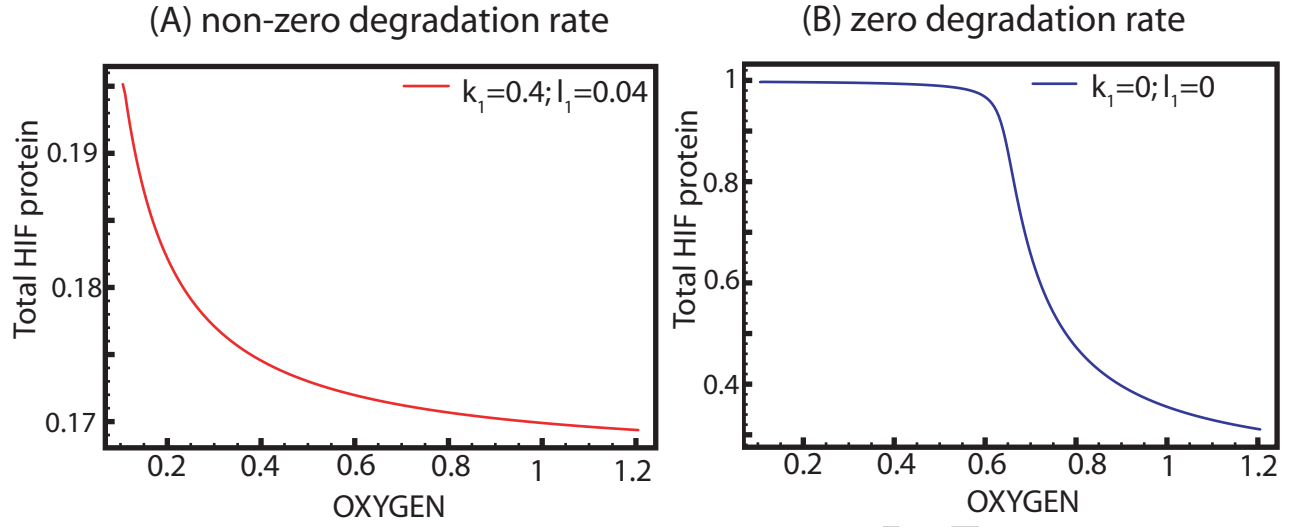


Figure 10:

$$\begin{aligned}
 (0) \quad \dot{x}_0 &= k_0 S_0 - l_0 x_0 - a_0 x_0 x_1 + d_0 x_8 - a_1 x_0 x_2 + d_1 x_9 - a_2 x_0 x_4 + d_2 x_{13} \\
 (1) \quad \dot{x}_1 &= k_1 S_1 - l_1 x_1 - a_0 (x_0 + x_6) x_1 + d_0 (x_8 + x_{11}) + b_0 [O_2]^2 (x_8 + x_{11}) \\
 (2) \quad \dot{x}_2 &= k_2 S_2 - l_2 x_2 - a_1 x_0 x_2 + d_1 x_9 + b_1 [O_2] x_9 \\
 (3) \quad \dot{x}_3 &= k_3 S_3 - l_3 x_3 - a_3 (x_5 + x_7) x_3 + (d_3 + b_3) (x_{10} + x_{12}) \\
 (4) \quad \dot{x}_4 &= -a_2 x_4 (x_0 + x_5) - a_4 x_4 (x_6 + x_7) + d_2 (x_{13} + x_{14}) + d_4 (x_{15} + x_{16}) \\
 (5) \quad \dot{x}_5 &= b_0 [O_2]^2 x_8 - a_2 x_4 x_5 + d_2 x_{14} - a_3 x_5 x_3 + d_3 x_{10} \\
 (6) \quad \dot{x}_6 &= -a_4 x_4 x_6 - l_0 x_6 + d_4 x_{15} + b_1 [O_2] x_9 - a_0 x_6 x_1 + d_0 x_{11} \\
 (7) \quad \dot{x}_7 &= b_0 [O_2]^2 x_{11} - a_4 x_4 x_7 + d_4 x_{16} - a_3 x_7 x_3 + d_3 x_{12} \\
 (8) \quad \dot{x}_8 &= a_0 x_0 x_1 - d_0 x_8 - b_0 [O_2]^2 x_8 \\
 (9) \quad \dot{x}_9 &= a_1 x_0 x_2 - d_1 x_9 - b_1 [O_2] x_9 \\
 (10) \quad \dot{x}_{10} &= a_3 x_5 x_3 - d_3 x_{10} - b_3 x_{10} \\
 (11) \quad \dot{x}_{11} &= a_0 x_6 x_1 - d_0 x_{11} - b_0 [O_2]^2 x_{11} \\
 (12) \quad \dot{x}_{12} &= a_3 x_7 x_3 - d_3 x_{12} - b_3 x_{12} \\
 (13) \quad \dot{x}_{13} &= a_2 x_0 x_4 - d_2 x_{13} \\
 (14) \quad \dot{x}_{14} &= a_2 x_4 x_5 - d_2 x_{14} \\
 (15) \quad \dot{x}_{15} &= a_4 x_6 x_4 - d_4 x_{15} \\
 (16) \quad \dot{x}_{16} &= a_4 x_4 x_7 - d_4 x_{16}
 \end{aligned}$$

Figure 11:

Tables

Table 1: Molecular species.

Species no.	Species identifier	Species no.	Species identifier
x_0	HIF	x_9	HIF:FIH
x_1	PHD	x_{10}	HIF(OH _p) ₂ : VHL
x_2	FIH	x_{11}	HIFOH _a :PHD
x_3	VHL	x_{12}	HIFOH _a (OH _p) ₂ :VHL
x_4	HRE	x_{13}	HIF:HRE
x_5	HIF(OH _p) ₂	x_{14}	HIF(OH _p) ₂ :HRE
x_6	HIFOH _a	x_{15}	HIFOH _a :HRE
x_7	HIFOH _a (OH _p) ₂	x_{16}	HIFOH _a (OH _p) ₂ :HRE
x_8	HIF:PHD		

Table 2: Set of enzymatic values consistent with numerical fitting. Enzymatic parameters units: t stands for time unit and $M = \text{Mol.l}^{-1}$ is the concentration unit.

A) Numerical values	B) Constraints	C) Enzymatic parameters		
$a = 0.36$	$\frac{k_0 S_0}{l_0} = \frac{a}{R}$	$k_0 = 0.0035 t^{-1}$	$l_0 = 1 t^{-1}$	$S_0 = 100 M$
$b = 11641.60$	$\frac{k_1 S_1}{l_1} = (\frac{\delta_1}{c} - 1) \frac{l_0}{a_1}$	$k_1 = 5.77 t^{-1}$	$l_1 = 1 t^{-1}$	$S_1 = 100 M$
$c = 26.42$	$\frac{k_2 S_2}{l_2} = \frac{\epsilon_1 l_0}{a_2}$	$k_2 = 635.34 t^{-1}$	$l_2 = 1 t^{-1}$	$S_2 = 100 M$
$\rho = 0.36$	$\frac{k_3 S_3}{l_3} = \frac{d_3 + b_3}{a_3 b_3} (\frac{\delta_1}{c} - 1) l_0$	$k_3 = 37.04 t^{-1}$	$l_3 = 1 t^{-1}$	$S_3 = 100 M$
$\delta_1 = 15269.20$	$\frac{d_1}{b_1} = \delta_1$	$a_1 = 1 t^{-1} M^{-1}$	$b_1 = 0.000065 t^{-1} M^{-2}$	$d_1 = 1 t^{-1}$
$\delta_2 = 6498.53$	$\frac{d_2}{b_2} = \delta_2$	$a_2 = 1 t^{-1} M^{-1}$	$b_2 = 0.000154 t^{-1} M^{-1}$	$d_2 = 1 t^{-1}$
$\epsilon_1 = 63534.3$	$a_4 = \frac{\rho l_0}{k_0 S_0} d_4$	$a_3 = 1 t^{-1} M^{-1}$	$b_3 = 1 t^{-1}$	$d_3 = 1 t^{-1}$
		$a_4 = 1.007 t^{-1} M^{-1}$		$d_4 = 0.1 t^{-1}$

Table 3: To measure the accuracy of the fitting, we computed the euclidean distance between the experimental and theoretical values. Namely: $d = ((156 - h(0.2))^2 + (153 - h(1))^2 + (117 - h(3))^2 + (41 - h(9))^2 + (9 - h(21))^2)^{\frac{1}{2}}$. The asterisk corresponds to the optimal value for n .

n	1	2	3	4	1.98*
d	33.38	2.35	13.90	19.83	2.29